

Kathrin Steyer / Meike Lauer / Annelen Brunner

Usuelle Wortverbindungen in OWID. Lexikografische Aspekte unter besonderer Berücksichtigung der Pragmatik

Abstract

Dieser Artikel fasst wichtige Aspekte der vom Projekt 'Usuelle Wortverbindungen' (UWV) erarbeiteten Konzeption für die korpusbasierte lexikografische Beschreibung von Wortverbindungen in OWID zusammen. Der Schwerpunkt in diesem Teilprojekt liegt auf der lexikografischen Beschreibung des typischen Gebrauchs von usuellen Wortverbindungen auf der Basis eines sehr großen Korpus des Deutschen. Zur differenzierten Untersuchung des Sprachgebrauchs werden korpusanalytische Methoden herangezogen und die Ergebnisse in einem nutzerfreundlichen Hypertextformat präsentiert. Zudem ist es ein Ziel, die sprachliche Vielfalt, die in den Korpora gerade auch in Bezug auf Wortverbindungen zu finden ist, durch eine große Menge authentischer Korpusbelege angemessen darzustellen.

This article presents important aspects of the corpus-based description of multi-word expressions in OWID, as developed by the project "Usuelle Wortverbindungen" (UWV). The main focus lies on the lexicographical description of the typical usage of multi-word expressions based on the evidence of a very large corpus of written German. Corpus analytical methods are employed to study language use and the results are presented in a user-friendly hypertext format. The diversity of language reflected in multi-word expressions is documented by a large number of authentic corpus samples.

Inhalt:

1. Die Erarbeitung der Musterartikel
2. Zur Mikrostruktur der Mehrwortartikel
 - 2.1 Lesartenübergreifende Angaben
 - 2.2 Lesartenspezifische Angaben
 - 2.3 Kommentare
3. Perspektiven
4. Literatur
5. Referenzwörterbücher

1. Die Erarbeitung der Musterartikel

In der ersten Arbeitsphase dieses Teilprojektes ging es um die Erarbeitung von Mehrwort-Musterartikeln, wobei insbesondere getestet werden sollte, welche Artikelstruktur sich in welchem Umfang für eine Mehrwortbeschreibung eignet. Des Weiteren sollte überprüft werden, mit welchem Arbeitsaufwand systematisches korpusbasiertes Vorgehen verbunden ist. Es wurden unterschiedliche Auswahlkriterien für die Lemmakandidaten in Abhängigkeit von der jeweiligen Zielsetzung zugrunde gelegt. Als Ergebnis entstanden zwei Artikeltypen: [25 Mehrwort-Musterartikel](#) und [100 Wortverbindungsartikel](#). Beide werden im Folgenden näher beschrieben.

Zur Ermittlung der Lemmakandidaten. Die Auswahl der Lemmakandidaten erfolgte durchweg auf der Basis der am IDS entwickelten [statistischen Kookkurrenzanalyse](#) (Bella 1995) mit unterschiedlichen Parametereinstellungen, derzeit u.a. implementiert in [COSMAS II](#).¹ Den Ausgangspunkt bildeten Kookkurrenzprofile von Einwortlemmata

¹ Die Artikelarbeit basiert auf dem Deutschen Referenzkorpus DeReKo, der mit über 3,2 Milliarden Wörtern weltweit größten linguistisch motivierten Sammlung elektronischer Korpora mit geschriebenen deutschsprachigen Texten aus der Gegenwart und der neueren Vergangenheit (<http://www.ids-mannheim.de/kl/projekte/korpora/>, Stand: 25.6.2008).

des [lexiko-Demonstrationswortschatzes](#) (alle Kookkurrenzpartner zum Nomen [Mobilität](#)), die vom [UWV-Projekt](#) linguistisch geprüft wurden.² Das obligatorische Kriterium für alle Lemmakandidaten stellte die statistisch auffällige Affinität zwischen den Basis-komponenten einer Wortverbindung dar. Diese Herangehensweise ist neuartig, da in der lexikografischen Praxis die Entscheidung für die Aufnahme in ein Wörterbuch stattdessen zumeist auf der Basis bereits vorhandener Idiomsammlungen und Nachschlagewerke getroffen wird. Dies führt in vielen Fällen zu einer Tradierung bestimmter phraseologischer Kernbestände, die nicht immer dem aktuellen Usus entsprechen. Mithilfe der statistischen Kookkurrenzanalyse kann man zwar nicht alle usuellen Wortverbindungen finden, die potenziell zum Bestand einer Sprachgemeinschaft gehören, da es viele gibt, die zu selten belegt und daher statistisch nicht auffällig sind. Allerdings lässt sich in einem Umkehrschluss sagen, dass diejenigen Wortverbindungen, die mithilfe der automatischen Methoden selektiert und als usuelle Wortverbindungen linguistisch qualifiziert wurden, eine gesicherte Festigkeit im Gebrauch aufweisen. Über die statistische Affinität hinaus mussten die ermittelten Kookkurrenzen eine Übersummativität aufweisen. Diese übersummativ Qualität ist jedoch nicht in einem engen Sinne als semantische Idiomatizität zu verstehen, sondern in einem weiten Sinne als ein „Mehr an Gebrauchsbedeutung“ des gesamten Mehrwortausdrucks im Vergleich zu den Bedeutungen ihrer Komponenten. Diese Wortverbindungen fungieren damit als kommunikative Einheiten.

25 Mehrwort-Musterartikel zu Demonstrationszwecken. Die Erarbeitung der 25 *Mehrwort-Musterartikel* (auf den OWID-Seiten eingeordnet unter [Mehrwortartikel vermisch](#)) sollte vor allem dazu dienen, die vom Projekt UWV entwickelte Artikelstruktur für die korpusbasierte Erarbeitung elektronischer Mehrwortartikel in der Praxis zu testen. Es wurden Wortverbindungen ausgewählt, die möglichst alle relevanten Mehrwortklassen, so wie sie die Mehrwortforschung bisher entwickelt hat (vgl. Burger 2007, Fleischer 1997), und alle in der Mehrwortartikelstruktur enthaltenen Angaben abdecken, z.B. verschiedene Idiomatizitätsgrade, Wortverbindungen mit mehreren Lesarten, verschiedene Mehrwortklassen wie Zwillingsformel oder Sprichwort, verschiedene Phrasentypen wie Präpositionalphrase oder satzwertige Wortverbindung oder verschiedene Realisierungen kommunikativer Funktionen wie Warnung oder Kritik (vgl. [verwendete Attributlisten](#)).³

100 Wortverbindungsartikel zu den Bezugsadjektiven *blind, geistig, gesund, normal, sanft*. In einem nächsten Arbeitsschritt wurden Wortverbindungen beschrieben, die sich aus der Auswertung der Kookkurrenzprofile ausgewählter Bezugswörter ergaben (auf den OWID-Portalseiten eingeordnet unter [Wortverbindungen zu](#)). Hier wurden die relevanten Informationen zum typischen Gebrauch im Korpus auf einige wenige Angaben konzentriert. Die Entscheidung für adjektivische Bezugswörter resultierte aus der Randständigkeit, die für diese Wortklasse in der bisherigen Mehrwortforschung noch festzustellen ist. Die Kookkurrenzanalyse ermöglicht die systematische Erfassung von usuellen Umgebungsmustern eines Adjektivs, die auf der Basis von Intuition nur schwer

² Zum Konzept der usuellen Wortverbindungen vgl. Steyer (2000), Steyer (i. Dr.). Zur linguistischen Interpretation und lexikografischen Anwendung von Korpusanalysemethoden vgl. u.a. Steyer (2004a; b), Steyer/Lauer (2007), Belica/Steyer (i.Dr.), Brunner/Steyer (2007). Zu aktuellen Tendenzen der phraseologischen Forschung und Phraseografie auch in Bezug auf korpusbasierte Empirie vgl. auch Hallsteinsdóttir/Farø (Hg.) (2006), Burger/ Dobrovol'skij/Kühn/Norrick (Hg.) (2007).

³ Zu einzelnen Aspekten der spezifischen Mikrostruktur dieses Artikeltyps verweisen wir auf den online verfügbaren [Informationstext der OWID-Portalseite 'Feste Wortverbindungen'](#).

generierbar sind. Ein Ergebnis dieses Teilprojektes war die Erkenntnis, dass auch viele transparent und zunächst nicht idiomatisch scheinende Wortverbindungen, vor allem bei Adjektiv-Nomen-Verbindungen (z.B. im Feld von *geistig* und *normal*), eine Festigkeit in der Verwendungsspezifik aufweisen und demzufolge als zusammenhängende Einheiten zu beschreiben sind. Dies führte zu einer primär pragmatisch motivierten Form von Gebrauchsbeschreibungen, die nicht immer identisch mit dem traditionellen Konzept von 'Bedeutungsparaphrasen' ist. Sie rückt die typischen kommunikativen Aspekte der jeweiligen Wortverbindung ins Zentrum.

2. Zur Mikrostruktur der Mehrwortartikel

2.1 Lesartenübergreifende Angaben

Artikelstichwort. Nach der Identifizierung einer Wortverbindung stellt die Entscheidung für das Artikelstichwort in einem Mehrwortartikel eine weitere schwierige lexikografische Aufgabe dar. Die vielfältigen Formen des Artikelstichworts für usuelle Wortverbindungen erfordern eine vorgelagerte Reflexion und vor allem systematische Korpusuntersuchungen (insbesondere KWIC-Analysen), um den festen Kern der Wortverbindung zu bestimmen. Dazu sind in den meisten Fällen mehrere Recherchevorgänge mit verschiedenen Parametereinstellungen notwendig, um wirklich alle Fälle in den Blick zu bekommen. Ein Syntagma kann dann als fester Kern gelten, wenn kein weiteres Element getilgt oder ersetzt werden kann, ohne damit die Bedeutung der Wortverbindung zu verändern. Zugleich sollte aber dieses Muster ohne weitere Ergänzungen verstehbar sein. Beispiele für solche Lemmazeichengestaltungsangaben sind:

- [*billiger Jakob*](#)
- [*Angst vor der eigenen Courage*](#)
- [*jenseits von Gut und Böse*](#)
- [*hoch und heilig*](#)
- [*an die große Glocke hängen*](#)
- [*im eigenen Saft schmoren*](#)
- [*auf Dauer*](#)
- [*Dienst nach Vorschrift*](#)

In vielen Fällen ist jedoch eine solche korrekte und zugleich leicht verständliche Benennung nicht realisierbar, da auch der semantische Kern vielfachen Variationen unterliegt. Hier müsste man eigentlich kategoriale Angaben verwenden. So ist etwa bei der Wortverbindung *da liegt der Hund begraben* (im Sinne von: 'dies ist der (tatsächliche) Grund / der Kern / die Ursache des Problems') nur das Partizip *begraben* fix (z.B. *da / hier / im Verwaltungshaushalt / in Nordhausen liegt / ist der Hund / Hase / die Krux begraben*). Die invariante Form lautet also: [DEIKT. ELEMENT / LOC + VERB + NOMEN + *begraben*]. Der Lexikograf steht vor einem Dilemma: Wenn das Artikelstichwort nur den möglichst unveränderbaren (obligatorischen) Kern einer Wortverbindung abbilden soll, kann das zu sehr abstrakten Darstellungen führen (siehe Beispiel). Dies steht aber im Widerspruch zu einer nutzerfreundlichen Form, die in einem lexikografischen Artikel unbedingt anzustreben ist.⁴ Deshalb wurde in den Artikeln in solchen Fällen das im Korpus häufigste Verwendungsmuster als Artikelstichwort angesetzt. Aus

⁴ Dobrovol'skij spricht in diesem Zusammenhang vom Widerspruch zwischen der lexikografischen Universalität der Nennform, die aber oft befremdlich wirke, und der psychologisch realen und damit benutzerfreundlicheren Aussageinformation, die aber nie alle Realisationsformen abdecken könne (Dobrovol'skij 1993, S. 59).

diesem Grund wurde etwa bei der Wortverbindung [eigen / eigenes Fleisch und Blut](#) die eigentlich ältere Form *eigen Fleisch und Blut* als Artikelstichwort angesetzt und bei der Wortverbindung [reden / sprechen, wie einem der Schnabel gewachsen ist](#) nur das Verb *reden* ins Artikelstichwort aufgenommen, da dieses am häufigsten in der Wortverbindung vorkommt. Derartige lexikografische Entscheidungen zu einem Artikelstichwort wurden jedoch immer mit einem Kommentar versehen, um deutlich zu machen, dass es sich nur um die typischste Form handelt. Darüber hinaus werden im Korpus vorgefundene Varianten der Wortverbindung, wenn sie eine auffällige Signifikanz aufweisen, gleichfalls angegeben.

Frequenzangabe (25 Mehrwort-Musterartikel). Die Angabe zur Vorkommenshäufigkeit basiert auf einer konkreten Suchanfrage, die die Wortverbindung in ihrem lexikalischen Kern abdeckt. Es ist zu beachten, dass die angegebene Zahl nur etwas über die Häufigkeit der Wortverbindungen auf der Basis dieser Suchanfrage und des zum Analysezeitpunkt zugrunde gelegten Korpus aussagt. Bei einer anderen Suchanfrage oder einer anderen Korpuszusammenstellung kann sich die Frequenz verändern. Ein Beispiel für diese Angabe findet sich im Artikel [an die große Glocke hängen](#): „Die Wortverbindung *an die große Glocke hängen* hat über 850 Vorkommen in allen IDS-Korpora der geschriebenen Sprache. (Stand: Oktober 2005)“.

Suchanfrage für Recherche im Korpus (100 Wortverbindungsartikel). Diese Angabe dient dazu, dem Nutzer den Zugang zu allen Vorkommen einer Wortverbindung auf dem jeweils aktuellsten Stand der öffentlich zugänglichen IDS-Korpora der geschriebenen Sprache zu ermöglichen, indem er die vorformulierten Suchanfragen nach dem Aufrufen von COSMAS II in das dortige Suchfenster kopiert. Die Formulierung von Suchanfragen für Wortverbindungen ist keineswegs trivial, da sie bereits auf Hypothesenbildungen basiert, das Beherrschen der Suchanfragesprache voraussetzt und oft auch ein gewisses Vorwissen in Bezug auf die Aussagekraft von Korpusanalysemethoden verlangt. Daher stellen Wortverbindungsanfragen an das Korpus nicht selten eine Barriere für Nutzer dar. In der Regel sind zwei Möglichkeiten – entsprechend kommentiert – angegeben: erstens eine Suchanfrage, mit der man relativ gesicherte Resultate erhält, die aber aufgrund ihres hohen Fokussierungsgrades nicht das ganze Spektrum der Variationen und Modifikationen erfassen muss, und zweitens eine Suchanfrage, die sehr viel weiter gefasst ist und daher diese produktiven Abwandlungsphänomene besser erfasst. Hier muss der Nutzer allerdings damit rechnen, dass er auch Resultate erhält, die überhaupt nicht die gesuchte Wortverbindung abbilden. Durch die dynamische Suche erhält der Nutzer darüber hinaus die Vorkommenshäufigkeit einer Wortverbindung auf dem aktuellsten Stand der öffentlich zugänglichen IDS-Korpora der geschriebenen Sprache. Ein Beispiel für diese Angabe findet sich im Artikel [im Vollbesitz seiner geistigen Kräfte](#): „(&Vollbesitz /+w2 &geistig) /+w1 &Kraft“.

Basiskomponente(n). Diese Angabe beinhaltet die Auflistung aller autosemantischen Komponenten einer Wortverbindung. Sie dient dazu, die Komponenten der Mehrwortartikel mit den jeweiligen Einwortlemmata im *lexiko*-Wörterbuch zu verbinden. So verweisen z.B. die Basiselemente im Artikel [kapitaler Hirsch](#) auf die *lexiko*-Artikel [kapital](#) und [Hirsch](#). Des Weiteren ist so ein Zugriff auf die beschriebenen Wortverbindungen über alle ihre Basiskomponenten möglich. So werden z.B. bei einer [Suche im gesamten OWID-Portal nach gesund](#) auch alle Wortverbindungsartikel, die die Basiskomponente *gesund* enthalten, geliefert.

Eine interne Annotation stellt die Angabe zur Unikalität dar, also zu solchen sprachlichen Einheiten, die keine freie Bedeutung außerhalb dieser Wortverbindung besitzen (z.B. das Element *frank* in der Wortverbindung *frank und frei*).

Variation(en) und Modifikation(en). Die Unterscheidung zwischen Variation und Modifikation basiert auf dem Konzept von Burger (2007). Unter Variation werden usuell Varianten einer Grundform verstanden, die das konzeptuelle Schema einer Wortverbindung nicht antasten (vgl. z.B. die Artikel *eigen Fleisch und Blut*, *reden, wie einem der Schnabel gewachsen ist*). Im Gegensatz dazu sind Modifikationen okkasionelle Abwandlungen einer Wortverbindung, die im Korpus sehr zahlreich zu finden sind. Diese Modifikationen wurden in einige Artikel testweise aufgenommen (vgl. z.B. *Tue Gutes und rede darüber*). Sie veranschaulichen den sprachlichen Reichtum, wie man ihn in den Korpora vorfindet, sehr gut. Es wird aber künftig noch zu überprüfen sein, welche Kriterien anzusetzen sind, um die lexikografische Relevanz von Variationen zu bewerten. Eine gewisse Musterhaftigkeit muss auch in den freieren Modifikationen erkennbar sein.

2.1 Lesartenspezifische Angaben

Gebrauchsbeschreibung. Eine korpusbasierte Beschreibung von Wortverbindungen ist in besonderer Weise an den usuellen Gebrauch gebunden. Wortverbindungen lassen sich als Mikrotexte mit Oberflächenstrukturen, propositionalem Gehalt und illokutiven Funktionen auffassen. Eine pragmatisch ausgerichtete Gebrauchsbeschreibung darf sich deshalb nicht auf eine eng semantische Paraphrase beschränken, sondern muss die Erfassung und Darstellung des kommunikativen Sinns einer Wortverbindung und ihrer Funktionen in der Kommunikation ebenso berücksichtigen. Wenn man beispielsweise nur die reine Bedeutung von *Dienst nach Vorschrift* mit 'seinen Beruf streng nach den Vorgaben ausüben' umschreiben würde, hätte man den kommunikativen Sinn und den typischen Gebrauch dieser Wortverbindung nicht erfasst. Erst durch den folgenden Kommentar wird der eigentliche Verwendungskontext deutlich: „Dabei kann sowohl gemeint sein, dass die Ausführung der Tätigkeit sehr genau geschieht, als auch, dass über die vereinbarten, vertraglich geregelten zu erbringenden Leistungen hinaus nicht mehr (zeitlich oder inhaltlich) gearbeitet wird. In der Regel wird *Dienst nach Vorschrift* als Druckmittel eingesetzt, oft anstelle eines Streiks.“ (vgl. *Dienst nach Vorschrift*).⁵ Das bedeutet, dass in sehr vielen Paraphrasen solche Formulierungen wie *Mit ... drücken Sprecher aus, dass / bewerten Sprecher ...* bevorzugt werden, wie etwa die Gebrauchsparaphrasen in den Artikeln *Kapital verspielen* und *recht und billig* zeigen.

Eine pragmatische Perspektive einzunehmen, bedeutet, folgende Fragen zu stellen:⁶

- Wer verwendet eine Wortverbindung wem gegenüber in welcher Situation mit welcher Intention?

⁵ Auch Burger diskutiert die unzureichende Aussagekraft einer Paraphrase, die auf den so genannten Bedeutungskern reduziert wird und nicht das damit einhergehende illokutive Potenzial berücksichtigt, u. a. anhand des usuellen Aufforderungsmodus des Phraseologismus „vor seiner eigenen Tür kehren“ (2007, S. 202f).

⁶ Es geht hier jedoch weniger um die in der phraseologischen Forschung relativ gut beschriebenen kommunikativen Formeln, Routineformeln usw., sondern vielmehr um sprachliche Einheiten, die nicht durch ein illokutives Potenzial direkt bestimmt sind. Einen guten Überblick über pragmatische Beschreibungsansätze in der Phraseologie bietet Filatkina (2007).

- Welche Verwendungsbedingungen und -beschränkungen gibt es für den Gebrauch einer Wortverbindung?
- Was geben Sprecher mit dem Gebrauch einer Wortverbindung über die eng an die Bedeutung gebundene Funktion hinaus noch zu verstehen?
- Worin besteht also der kommunikative Sinn einer Wortverbindung?

Die wichtigsten methodischen Zugänge sind heuristische Kookkurrenzfelder, KWIC- und Beleganalysen, wie am Ausschnitt aus der Umgebungsanalyse von *billiger Jakob* gut sichtbar wird:

Total	Anzahl	Autofokus von bis	LLR	Kookkurrenzen	syntagmatische Muster
5	5	1 1	524	Jakobs Image	100% Image des billigen Jakobs
27	22	1 1		Jakobs	54% des die billigen Jakobs
35	8	-2 -2	188	nicht spielen	75% nicht ... den billigen Jakob ... SPÖ spielen
46	11	-2 -2		nicht sind	54% Wir sind [...] nicht der billige Jakob aber
55	9	-2 -2		nicht werden	44% nicht zum billigen Jakob zu werden
93	38	-2 -2		nicht	28% das Sojus nicht als den billigen Jakob an
96	3	-2 -2	96	Image	33% Wir ... Image ... Billigen
101	5	2 5	76	spielen	80% den billigen Jakob zu spielen

In den entsprechenden KWICs werden Muster deutlich, die Aufschluss über eine mit dieser Wortverbindung einhergehende negative Konnotation geben:

Jakobs Image:

L99 erwartet BM/dpa/StM Frankfurt/M. - Praktiker will das Image des "**billigen Jakobs**" unter den Baumarkt-Ketten loswerden. Der neue
 E97 Um vom auf die Dauer ohnehin unvorteilhaften Image des **billigen Jakobs** wegzukommen, besinnt sich das Fürstentum auf seine
 E98 Sortiment, die anderswo schwer erhältlich sind. Das Image eines **billigen Jakobs** für Wald-und-Wiesen-Bands oder zweitrangige Sänger und
 R99 radikal umkrempeln. Vor allem soll die Metro-Tochter das Image des "**billigen Jakobs**" loswerden. "Dauerhaft kann die Kundenbindung nur
 P98 führt er Titel, die schwer erhältlich sind. Das Image des **billigen Jakobs** für Wald-und-Wiesen-Bands oder zweitrangige Sänger läßt

nicht spielen:

B99 Vorstandschef des Fernsprechriesen verkündet, künftig nicht den "**billigen Jakob**" spielen zu wollen. Davon ist die Telekom nun abgerückt.
 T91 so groß geworden zu sein, daß sie sich auch nicht mehr scheut, den **billigen Jakob** zu spielen. So wird die soziale Fallstudie als ein »Stück
 I99 gereizt. Er werde nach der Nationalratswahl sicher nicht den "**billigen Jakob**" spielen, vertraute Schüssel dem Wiener Standard an: Es
 A98 Gebührenerhebung beim Postkonto. «Wir können nicht länger den **billigen Jakob** spielen und Top-Leistungen anbieten, ohne etwas dabei zu
 I99 gestellt wurden. Schüssel wolle zwar nicht den "**billigen Jakob**" gegenüber der SPÖ spielen, wohl aber "der billige
 R99 Wettbewerber schlichtweg unterschätzt. "Wir wollen nicht den **billigen Jakob** spielen", lautete sein Credo. Mittlerweile ist das
 M98 mit insgesamt 20 000 Mitarbeitern dürften nicht länger den "**billigen Jakob**" spielen, warnte Hariegel. Sie müßten mehr auf Qualität,
 K99 Koalition. Der VP-Obmann hatte gemeint, man wolle nicht weiter den "**billigen Jakob** für die SPÖ" spielen. Ursula Stenzel billigte Rudas in

nicht sind:

P94 Angebot, kommen aber mit keinen Schleuderpreisen. Der **billige Jakob** sind wir nicht, sondern wollen Kunden möglichst viel Auto
 HMP05 nicht allein in punkto Preis absetzen: "Wir sind nicht der **billige Jakob**, sondern bieten die beste Mischung aus Preis und
 T93 Kröte zu schlucken. Und fügt trotzig hinzu: "Wir sind nicht der **billige Jakob**." Genau den glauben manche Sozialdemokraten in Markus
 T89 der SV Waldhof Mannheim, fühlte sich unwohl: "Wir sind nicht der **billige Jakob**, aber wir können wohl nichts dagegen machen", erklärte
 T98 müßten auch unangenehme Wahrheiten offen gesagt werden. "**Billiger Jakob** sind die anderen, nicht wir." Die Union habe die bessere

T01 des Angebots seines Angestellten reiben: "Wir sind zwar nicht der **billige Jakob**, aber Qualität macht sich eben bezahlt."
 B98 sein, es ist nicht so bequem mit der Union, nicht so billig. **Billiger Jakob** sind die anderen, nicht wir. So wird man bei jedem Thema
 I98 keine "Wegelagerer" am Brenner seien "wir sind aber auch nicht der **billige Jakob** im Transitverkehr". Der Tiroler SPÖ-Mandatar Erwin
 N98 wo er die Linie "Habsburg" mitentwickelte. "Wir sind nicht der **billige Jakob**, aber wir haben eben andere Kalkulationsvorstellungen als
 P93 Kasse." Denn, darüber sind sich Mayr und Seitlinger einig: "Der **billige Jakob** sind wir nicht. "
 B00 Vorstellungen zum Kaufpreis". "Wir sind nicht der **billige Jakob**", sagt er. Schließlich habe sich die Mega an die WoGeHe

nicht werden:

T02 Die Stiftung warnt, Engagement dürfe nicht als „**billiger Jakob** des Sozialstaats“ missbraucht werden. Was für
 I97 der Tiroler VP-NR Dieter Lukesch. Denn Tirol dürfe nicht zum **billigen Jakob** des Transitverkehrs werden. Konkret verlangte er
 R98 und wettbewerbsgerechte Preise" bieten, dabei aber nicht zum "**billigen Jakob**" werden. In anderen Ländern habe sich gezeigt, daß die
 Z98 Telekom-Chef Ron Sommer sich fest vorgenommen hatte, nicht zum "**billigen Jakob**" zu werden, ändert er nun abrupt seinen Kurs: Mit einer
 Z02 Auslandsgesprächen. Und obwohl Sommer angekündigt hatte, nicht zum "**billigen Jakob**" zu werden, blieb ihm schließlich doch nichts anderes,
 DPA06 sinken. Ricke beteuerte zugleich, dass der Konzern nicht der «**billige Jakob**» werden wolle. Wichtig seien auch Qualität und Service.
 DPA06 Ricke an. Zugleich beteuerte er, dass der Konzern nicht der «**billige Jakob**» werden wolle. Wichtig seien auch Qualität und Service.
 DPA06 Trotz der Preisabschläge wolle der Bonner Konzern nicht der «**billige Jakob**» werden, betont Ricke. «Den Kampf um die Kunden werden
 F97 Die Werkstätten für Behinderte dürften nicht als der "**billige Jakob** des Arbeitsmarkts" betrachtet werden. Sie stünden unter

Diese Erkenntnisse können lexikografisch umgesetzt werden – z.B. in Gebrauchsbeschreibungen wie im Artikel [billiger Jakob](#):

Mit *billiger Jakob* wird jemand bezeichnet, der ein Produkt oder eine Leistung zu niedrigen Preisen, oft unter Verzicht auf Anspruch und Qualität, anbietet. *Billiger Jakob* wird auch im Sinne eines 'Sich-unter-Wert-Verkaufens' oder eines 'nützlichen Idioten' verwendet. In den meisten Fällen ist damit eine negative Bewertung verbunden.

Wertungsaspekt: Mit *billiger Jakob* verbinden Sprecher eine negative Bewertung, die sich zum Teil darin ausdrückt, dass eine solche Zuschreibung zurückgewiesen wird.

Gerade in Bezug auf den pragmatischen Mehrwert, der mit einer sprachlichen Einheit verbunden sein kann, ermöglicht die Korpusanalyse gesicherte Aussagen, die über die individuelle Einzelwahrnehmung eines Sprechers hinausgehen. In der Lexikologie und Lexikografie wird u.E. bisher zu wenig deutlich gemacht, dass auch funktionale Aspekte einer abgestuften Festigkeit unterliegen, was unter Umständen auch zu Überinterpretationen und unzulässigen Generalisierungen führen kann. Deshalb unterscheiden wir zunächst prinzipiell zwischen zwei Perspektiven:

- a) verwendungsspezifische Komponenten, die konstitutiv für die Bedeutung einer Wortverbindung sind (**bedeutungsinhärente Gebrauchsspezifik**);
- b) verwendungsspezifische Komponenten, die eine Wortverbindung kontextuell und situativ spezifizieren bzw. ausdifferenzieren (**situative Gebrauchsspezifik**).⁷

⁷ Zentrale Aspekte unseres Ansatzes für pragmatisch motivierte lexikografische Angaben sind auch in die Artikelstruktur des *lexiko*-Wörterbuchs für die Einzelwortbeschreibung eingeflossen (vgl. dazu auch Haß (Hg.) 2005, S. 265).

Bedeutungsinhärente Gebrauchsspezifik meint die Gebundenheit der Bedeutung einer Wortverbindung an eine oder mehrere ganz bestimmte Gebrauchsbedingungen, also ihre funktionale Festigkeit. Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf diese Gebrauchsbedingungen einzugehen. Bedeutungsinhärente Gebrauchsspezifik kann alle Aspekte der Verwendung betreffen und stellt ein obligatorisches Merkmal dar, wie folgende Beispiele zeigen:

Kein Grund zur Panik!: Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf den damit einhergehenden Sprechakt der Beruhigung einzugehen.

geistiger Brandstifter: Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf die negative Konnotation einzugehen.

hoch und heilig: Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf die emotionalisierende Verstärkungsfunktion der Zwillingsformel in Bezug auf ein Verb wie *versprechen* einzugehen.

eingeschränktes Gesichtsfeld: Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf die Domänengebundenheit bzw. die Fachsprachlichkeit dieses Ausdrucks (Augenheilkunde) einzugehen.

sanft entschlafen: Man kann die Bedeutung der Wortverbindung nicht erklären, ohne auf die Textsortengebundenheit (Todesanzeigen, Grabsteininschriften) einzugehen.

Somit können potenziell alle Aspekte der Verwendung einer Wortverbindung eine zentrale Stellung für die Erklärung der Bedeutung einnehmen. In unseren Artikeln betrifft dies vor allem: den Wertungsaspekt, die Funktion in der Kommunikation, das Sachgebiet, die Gruppensprache, die Textspezifik, die Medialität.

Die kontextuell-situative Gebrauchsspezifik betrifft demgegenüber typische Einstellungen, Situationen usw., die zur Bedeutung einer Wortverbindung hinzutreten und die durch die Korpusanalyse empirisch identifiziert werden können. Diese sind aber nicht obligatorisch für den Bedeutungskern. Bei diesen situativen Faktoren, die den Gebrauch einer Wortverbindung kontextuell spezifizieren, geht es nicht um das Sprachverhalten individueller Sprecher (etwa in einer aktuellen Face-to-Face-Situation) oder um konkrete Interaktionen (Gespräch, Einzeltext usw.). Es geht vielmehr um die Verwendungsmuster kollektiver Sprecher, deren Sprachverhalten sich in den Korpora manifestiert. Damit sind für uns auch hier nur jene Phänomene von Interesse, denen man aufgrund ihrer Rekurrenz im Korpus Musterhaftigkeit zusprechen kann.

In den 25 Musterartikeln wurden diese zwei verschiedenen verwendungsspezifischen Aspekte intern noch unterschiedlich ausgezeichnet und in verschiedenen lexikografischen Angaben verarbeitet. In den 100 Wortverbindungsartikeln floss diese Unterscheidung in die entsprechende Formulierung der Gebrauchsbeschreibung ein. Situative Aspekte der Verwendungsspezifik werden in den Artikeln durch einschränkende Adverbien wie *häufig*, *auch*, *überwiegend* oder *zumeist* oder auch typografisch durch einen neuen Absatz markiert. Diese Herangehensweise illustrieren die folgenden drei Beispiele:

Angst vor der eigenen Courage

Mit *Angst vor der eigenen Courage* bewerten Sprecher Absichten, Handlungen bzw. Folgen von Handlungen **zumeist** negativ.

geistige Höhenflüge

In den IDS-Korpora der geschriebenen Sprache werden *geistige Höhenflüge* auf **überwiegend** ironische Weise thematisiert, entweder um geistige Leistungen anderer gerade als unterdurchschnitt-

lich zu bewerten oder fremde Erwartungshaltungen gegenüber der eigenen Leistung zu bremsen und dem Adressaten zu vermitteln, dass er keine übertriebenen Erwartungshaltungen hegen dürfe.

der ganz normale Wahnsinn

Mit *der ganz normale Wahnsinn* werden Zustände oder Situationen, die extrem oder zugespitzt zu sein scheinen, doch als alltäglich bzw. völlig normal bewertet.

Häufig soll mit dieser Bewertung auch eine gewisse Gelassenheit ausgedrückt werden, verbunden mit der Ermunterung, die Dinge soweit möglich hinzunehmen, wie sie eben mal sind.

Auch die Beschreibung der situativen Faktoren des typischen Gebrauchs einer Wortverbindung kann sich auf sehr verschiedene Aspekte der Verwendungsspezifika beziehen.

Zur Erfassung und Beschreibung der situativen Gebrauchsspezifika:

Wertungsaspekt.⁸ Hier geht es um die Erfassung und Beschreibung typischer Einstellungsmuster, die Sprecher mit der Wortverbindung noch transportieren und die durch die Korpusanalyse als rekurrente Muster erkennbar werden, z.B.

geistiger Vater: Die Wortverbindung ist im Kern ohne auffällige Konnotation. Erst durch bestimmte Verwendungskontexte wird ihr eine negative Bewertung situativ zugeschrieben, z.B. wenn die Idee / Bewegung / Richtung, für die jemand als Urheber gilt, im aktuellen Diskurs negativ bewertet wird. Dies lässt sich anhand von typischen Kontextmustern wie *geistiger Vater des Fundamentalismus, des Krieges, von Auschwitz, der Terroristen* aus den Korpusbefunden identifizieren.

geistige Frische: Dem an sich schon positiv bewerteten Sachverhalt, dass jemand noch gute geistige Fähigkeiten aufweist, wird in bestimmten Verwendungskontexten eine weitere Bewertung im Sinne von Erstaunen oder Bewundern hinzugefügt: *in erstaunlicher / beneidenswerter geistiger Frische*.

gesundes Selbstbewusstsein: Mit der Wortverbindung verbinden Sprecher häufig eine Bewertung, die aber in der Regel vage gelassen wird. Die Korpusbefunde legen sowohl negative als auch positive Bewertungen des Sachverhalts, dass jemand ausgeprägt selbstbewusst ist, nahe. Die Interpretation hängt dann letztlich vom jeweiligen Wertekanon des Rezipienten ab.

Funktion in der Kommunikation. Sprecher vollziehen mit Wortverbindungen oft gerichtete Sprachhandlungen, z.B. einen Befehl, eine Bitte, eine Drohung, eine Warnung. Beispiele für solche in Wortverbindungen realisierten kommunikativen Funktionen sind: Kapital verspielen (wird oft als Warnung formuliert: *darf nicht verspielt werden*); Tue Gutes und rede darüber (ironische Zurückweisung / Kritik).

Sachgebiet, Gruppensprache, Textspezifika, Medialität. Diese Informationen sind bezogen auf auffällige Gebundenheiten im Gebrauch der Wortverbindung in Bezug auf die kommunikative Situation im weitesten Sinne. Sie können aus den Korpusbefunden natürlich nur vermittelt generiert werden, da alle diese Merkmale über schriftliche, in den IDS-Korpora vor allem massenmedial verbreitete Texte transportiert werden. Indikatoren für derartige Gebrauchsspezifika sind jedoch eine auffällig häufige metasprachliche Kommentierung, z.B. zum fachsprachlichen Status oder zur eigentlichen fach-

⁸ Zu linguistischen Aspekten von 'Bewertung' und entsprechenden sprachlichen Realisierungen vgl. Fries (1991).

sprachlichen Herkunft einer Wortverbindung, oder auffällig häufige Markierungen von Mündlichkeit. Informationen zum Sachgebiet beziehen sich auf ein auffälliges Vorkommen in einer oder mehreren Domänen im Korpus, wie beispielsweise Wirtschaft, Sport, Finanzwesen oder Politik, z.B. [blinder Fleck](#) (Augenheilkunde); [gesunde Härte](#) (Sport).

Die gruppensprachliche Information bezieht sich auf eine im Korpus auffällige Thematisierung von Vorkommen einer Wortverbindung im Sprachgebrauch bestimmter Sprechergruppen, z.B. [kapitaler Hirsch](#) (Jägersprache).

Die textsortenbezogenen Informationen dienen zur Markierung der typischen textsorten- bzw. typengebundenen Verwendung einer Wortverbindung. Eine Wortverbindung kommt im Korpus auffällig häufig in einer oder mehreren Textsorten vor, z.B. [geistige Frische](#) (Geburtstagsanzeigen / Glückwünsche).

Medialität lässt sich in schriftlichen Korpora nur insoweit erfassen, als dass die Vorkommen im Korpus auf auffallend häufige Markierungen von Mündlichkeit überprüft werden können (z.B. Anführungszeichen, Interjektionen) oder die Zugehörigkeit zu einer bestimmten medialen Kommunikationsform sehr explizit und wiederholt metasprachlich kommentiert wird, z.B. [noch ganz gesund sein](#); [Kein Grund zur Panik!](#).

Typische Kontextmuster. Die Kontextmusterangabe dokumentiert die typischen, durch automatische Umgebungsanalysen der jeweiligen Phrase gewonnenen, syntagmatischen Einbettungen der Wortverbindung. Die Kontextmuster enthalten keine systematischen Aussagen zur Bildbarkeit von Wortverbindungen, also präskriptive Angaben von Phrasenstrukturen oder Restriktionen (z.B. in Form von regelhaften Beispielphrasen). Das Spektrum typischer kontextueller Muster, so wie sie in der natürlichsprachigen Umgebung verwendet werden, dient vielmehr als Illustration des typischen Gebrauchs der Wortverbindung im Korpus. Die einzige Abstraktion wird dadurch getroffen, dass im Regelfall die übliche Nennform (z.B. bei Nomen: Nominativ Singular) angesetzt wird, es sein denn, es gibt Restriktionen innerhalb der Flexionsparadigmen oder Satzpositionen (also z.B. nur Pluralverwendungen, nur Genitiv, nur partizipial, nur prädikativ, nur nebengeordnet usw.). Diese Orientierung an den Resultaten der automatischen Umgebungsanalyse bedeutet in manchen Fällen, dass die Kontextmuster gemessen an regelhaften grammatischen Phrasen oder Sätzen unvollständig erscheinen. Damit soll dokumentiert werden, dass genau diese sprachlichen Komponenten als eine feste sprachliche Einheit erkannt und selektiert wurden. Zu jedem Kontextmuster werden Belege zur Illustration angegeben, sodass man die Kontextmuster quasi als Label für die eigentlichen Volltextstellen verstehen kann. Die Artikelstruktur bietet die Möglichkeit, Kontextmuster in Gruppen zusammenzufassen. Auch hier können die Kriterien für eine solche Kontextmustergruppe alle Ebenen des Sprachsystems betreffen; sie können aber auch andere Kriterien des Gebrauchs folgen, z.B. thematisch-diskursiven Zusammenhängen.

[Kapital verspielen](#)

ein / ihr / ... Kapital verspielen

viel Kapital verspielen

Kapital darf nicht verspielt werden

Gefahr, Kapital zu verspielen

politisches Kapital verspielen

wichtiges Kapital verspielen

wertvolles Kapital verspielen

Kapital leichtsinnig verspielen

Kapital leichtfertig verspielen

Kapital leicht verspielen

Beziehungen zu anderen Wörtern und Wortverbindungen. In dieser Angabe werden Verwendungszusammenhänge zwischen Wortverbindungen oder Beziehungen im Hinblick auf Wortklassen thematisiert, die durch vergleichende Kookkurrenzanalysen hervorgetreten sind. Diese Verwendungszusammenhänge lassen sich in den meisten Fällen nicht mit tradierten Kategorien wie etwa den paradigmatischen Relationen fassen. So gibt es aus korpusbasierter Sicht eigentlich nie eine echte Synonymie. Durch die Möglichkeit der feinen Kontextspezifikation treten auch distinktive Eigenschaften und Gebrauchsbesonderheiten einer Wortverbindung im Kontrast zu einer anderen deutlich zutage.

Folgende Beispielartikel enthalten Verweise zu anderen Wortverbindungen, die das Spektrum solcher Wortverbindungsfelder wenigstens andeuten:

[*blind gegenüber*](#) – [*auf einem Auge blind sein*](#) – [*sich blind und taub stellen*](#) – [...] [*macht blind*](#) – [*Liebe macht blind*](#)

[*geistiger Brandstifter*](#) – [*geistiger Vater*](#) – [*geistiger Wegbereiter*](#)

2.3 Kommentare

Neben den einzelnen lexikografischen Angaben nehmen auch frei formulierte Kommentare einen wichtigen Platz ein. Diese („Hintergrund“-)Kommentare können sowohl spezifische Aspekte des Gebrauchs betreffen als auch Fachkommentare darstellen, etwa zu einer speziellen lexikografischen Entscheidung, zum Gebrauch einzelner Basiskomponenten oder zu Zusammenhängen mit Inhalten anderer Wörterbücher (in OWID oder extern). Eine feinere Unterteilung, beispielsweise in Fachkommentare und Kommentare zum Gebrauch, wie sie noch bei den 25 Musterartikeln vorgenommen wurde, hat sich nicht bewährt und wurde für die 100 Artikel verworfen. Beispiele für Kommentierungen finden sich in allen Teilen der Mikrostruktur:

Kommentare zu „Artikelstichwort“

Mehrwortstatus:

[*Blind Date*](#): *Blind Date* wurde auf Grund des signifikanten Kovorkommens als Stichwort angesetzt: *Date* ist ein statistisch auffälliger Kookkurrenzpartner von *blind*. Im [*Neologismenwörterbuch \(NWB\)*](#) wurde '[*Blind Date*](#)' dagegen als Einwortlemma angesetzt, da dort auch Anglizismen in Form von Wortgruppen als Einzelwörter behandelt werden.

Schreibung:

[*Blind Date*](#): *Blind Date* ist mit ca. 500 Treffern die häufigste Schreibvariante in den IDS-Korpora der geschriebenen Sprache. Weniger häufig sind die Schreibvarianten *blind date* (35 Treffer) und *Blind date* (22 Treffer).

Flexionsform:

die normalste Sache der Welt: Die Wortverbindung ist in den IDS-Korpora der geschriebenen Sprache ausschließlich im Superlativ belegt.

Lexikalische Festigkeit:

die sanfte Tour: Das Artikelstichwort wurde mit dem bestimmten Artikel *die* angesetzt, da der bestimmte Artikel usuell ist.

Typik:

im Vollbesitz seiner geistigen Kräfte: Die Wortverbindung *im Vollbesitz der / seiner geistigen Kräfte* ist die signifikanteste Realisierung des Grundmusters *im Vollbesitz von etw. sein*. Andere Ergänzungen sind beispielsweise *seiner Fähigkeiten, ihrer Macht, der Wahrheiten, ihrer Stimme*.

Kommentare zu „Variation / Modifikation“

auf Dauer: Die Wortverbindung *auf Dauer* weist ein Vorkommen in den IDS-Korpora geschriebener Sprache von über 20 000 Treffern, die Nebenform *auf die Dauer* von über 4000 Treffern auf.

Gut gebrüllt, Löwe: Die Wortverbindung *Gut gebrüllt, Löwe* wird gelegentlich modifiziert, indem statt des Lexems *Löwe* Namen oder Bezeichnungen von Personen, Vereinen, Institutionen oder Ähnlichem verwendet werden.

sanfte Revolution: Variation: *samtene Revolution*: Die Wortverbindung *samtene Revolution* ist in den IDS-Korpora der geschriebenen Sprache häufiger belegt. Sie kann allerdings nur als teilsynonym gelten, da sie im Korpus ausschließlich als Bezeichnung für den politischen Umbruch in der Tschechoslowakei 1989/1990 verwendet wird.

Kommentare zu „Suchanfrage für Recherche im Korpus“

In einem gesunden Körper wohnt ein gesunder Geist: (&gesund /w1 &Körper) /s0 &Geist: Diese Suchanfrage bietet das genaueste Ergebnis. Wenn Sie an einem möglichst großen Modifikationsspektrum interessiert sind, sollten Sie die lexikalischen Einheiten innerhalb eines Satzes mit folgender Anfrage suchen: (&gesund /s0 &Körper) /s0 &Geist. Sie müssen dann aber damit rechnen, dass in bestimmten Fällen auch Kontextstellen angezeigt werden, die nicht zum Verwendungsspektrum dieser Wortverbindung gehören.

Kommentare zu „Gebrauchsbeschreibung“

auf Dauer (Lesart 'auf lange Sicht'): *Auf Dauer* wird in dieser Lesart sehr häufig in Negationskontexten verwendet. Damit wird ausgedrückt, dass einer weiteren Zukunft eines Sachverhalts oder eines Zustands keine Chancen eingeräumt werden bzw. diese Zukunftsperspektive nicht für sinnvoll erachtet wird.

blinder Aktionismus: (Mit *blinder Aktionismus* werden Entscheidungen oder Handlungen als übertrieben, voreilig, wenig durchdacht und damit als wenig hilfreich, wenn nicht gar als kontraproduktiv bewertet ...)

Kommentar: Kookkurrenzpartner von *Aktionismus* wie *hektisch, wild, blank, sinnlos, blindwütig* oder *kurzfristig* stützen diese Gebrauchsbeschreibung.

geistiger Brandstifter: Die Wortverbindung *geistiger Brandstifter* ist in den IDS-Korpora der geschriebenen Sprache seit Beginn der 90er Jahre belegt.

Kommentare zu „Kontextmuster“

blinder Alarm: Kontextmuster *sich als blinder Alarm herausstellen*: Die ergänzenden Verben dieser Gruppe sind an die Wortverbindung *blinder Alarm* gebunden und kommen nie oder nur selten allein mit dem Nomen *Alarm* kombiniert vor.

Kommentare zu „Beziehungen zu anderen Wörtern und Wortverbindungen“

geistiger Vater: Die Wortverbindung steht in einem Verwendungszusammenhang mit geistiger Wegbereiter; geistiger Brandstifter.

Kommentar: Die Wortverbindungen *geistiger Wegbereiter* und *geistiger Brandstifter* werden in den IDS-Korpora der geschriebenen Sprache jedoch überwiegend (*Wegbereiter*) oder immer (*Brandstifter*) negativ konnotiert verwendet, was an signifikanten Verwendungsmustern wie *geistiger Wegbereiter des Nationalsozialismus*; *geistiger Wegbereiter des Faschismus*; *geistiger Wegbereiter des Terrorismus* erkennbar wird.

Wer zu spät kommt, den bestraft das Leben: Bei der Wortverbindung *Wer zu spät kommt, den bestraft das Leben* handelt es sich um eine Wortverbindung, die die strukturellen und funktionalen Eigenschaften eines Sprichworts bereits erfüllt, obwohl auch heute noch die Autorschaft und der ursprüngliche Äußerungszusammenhang präsent sind. Daher ist sie sowohl als Sprichwort als auch als Zitat einzuordnen.

Kommentare in Verbindung zu anderen lexikografischen Ergebnissen / externen Wörterbüchern / Internetquellen⁹

jenseits von Gut und Böse: Die Wortverbindung *jenseits von Gut und Böse* geht auf Friedrich Nietzsches philosophisches Werk „Jenseits von Gut und Böse. Vorspiel einer Philosophie der Zukunft“ von 1886 zurück. Die Autorschaft dieser Wortverbindung ist in den aktuellen Korpusbelegen nach wie vor sehr präsent, sodass man diese Wortverbindung sowohl als Phraseologismus im engeren Sinne als auch als Zitat einordnen kann.

3. Perspektiven

Die lexikografische Beschreibung usueller Wortverbindungen wird in den nächsten Jahren mit der Erarbeitung von Mehrwortartikeln zu ausgewählten Feldern weitergeführt. Die entsprechenden Annotationen sollen der späteren Recherchierbarkeit nach linguistischen Merkmalen dienen und, wo es möglich ist, analog zu den entsprechenden Attributen im *ellexiko-Wörterbuch* angesetzt werden. Der Zugang zu den Artikeln wird stets auf zweifache Weise möglich sein: zum einen über die projektspezifische Stichwortliste auf unserer *OWID-Portalseite*, zum anderen über die Stichwortsuche im gesamten OWID-Portal. Die *UWV-Stichwortliste* unterscheidet sich von einer herkömmlichen, z.B. der *300 000er Stichwortliste* des *ellexiko-Wörterbuchs*, dadurch, dass es keine im Voraus definierte Lemmaliste geben wird, sondern dass sie erst durch die Auswertung systematischer Kookkurrenzanalysen während der laufenden Forschungsarbeit des Projektes ‘Usuelle Wortverbindungen’ sukzessive entsteht. Vorgesehen ist, die OWID-Artikel mit anderen Projektergebnissen, die in elektronischer Form präsentiert werden, zu verlinken. Vorstellbar ist darüber hinaus, mit externen Partnern erarbeitete lexikografische Daten (Artikel, korpusvalidierte Stichwortlisten zu einzelnen Mehrwortklassen) zu integrieren.

⁹ Auch diese Kommentare beziehen sich auf Beobachtungen des typischen Gebrauchs im Korpus. Es werden also nicht systematisch externe Quellen analysiert oder entsprechende Einträge mit unseren Informationen verglichen. Diese Verweise kommen nur dann ins Spiel, wenn sich in den Korpora auffällig häufige Indizien z.B. für ein Zitat oder eine spezifische Herkunft oder auch für enzyklopädische Informationen finden. Im Beispiel *jenseits von Gut und Böse* wird die Quelle (F. Nietzsche) häufig in den Korpusbelegen angeführt. Sie ist quasi in den Köpfen der Sprecher noch präsent (vgl. auch *Wer zu spät kommt, den bestraft das Leben*).

4. Literatur

- Belica, Cyril (1995): Statistische Kollokationsanalyse und Clustering. Korpuslinguistische Analysemethode. Mannheim. Internet: <https://cosmas2.ids-mannheim.de/cosmas2-web/> (Stand: Mai 2008).
- Belica, Cyril/Steyer, Kathrin (i. Dr.): Korpusanalytische Zugänge zu sprachlichem Usus. In: AUC (Acta Universitatis Carolinae), Germanistica Pragensia XX. Prag 2006. [Vorabdruck im Internet unter http://www.ids-mannheim.de/ll/uvw/dateien/Belica_Steyer_2005.pdf] (Stand: Mai 2008).
- Brunner, Annelen/Steyer, Kathrin (2007): Phraseologische und phraseographische Aspekte korpusgesteuerter Empirie. In: Jesenšek, Vida/Fabčič, Melanija (Hg.): Phraseologie kontrastiv und didaktisch. Neue Ansätze in der Fremdsprachenvermittlung. (= Zora 47). Maribor, S.181-194.
- Burger, Harald (2007): Phraseologie. Eine Einführung am Beispiel des Deutschen. 3., neu bearb. Aufl. (= Grundlagen der Germanistik 36). Berlin.
- Burger, Harald/Dobrovol'skij, Dmitrij/Kühn, Peter/Norrick, Neal R. (Hg.) (2007): Phraseologie / Phraseology. Ein internationales Handbuch zeitgenössischer Forschung / An international handbook of contemporary research. 2 Halbbde. (= HSK 28.1/2). Berlin/New York.
- Dobrovol'skij, Dmitrij (1993): Datenbank deutscher Idiome. Aufbauprinzipien und Einsatzmöglichkeiten. In: Földes, Czaba (Hg.): Germanistik und Deutschlehrerausbildung. Szeged/Wien, S. 51-67.
- Fleischer, Wolfgang (1997): Phraseologie der deutschen Gegenwartssprache. 2., durchges. u. erg. Aufl. Tübingen.
- Filatkina, Natalia (2007): Pragmatische Beschreibungsansätze. In: Burger/Dobrovol'skij/Kühn/Norrick (Hg.), Teilbd. 1, S. 132-158.
- Fries, Norbert (1991): Bewertung. Linguistische und konzeptuelle Aspekte des Phänomens. In: Sprache und Pragmatik. (= Arbeitsberichte 23). Lund, S.1-31.
- Hallsteinsdóttir, Erla/Ken Farø (Hg.) (2006): Neue theoretische und methodische Ansätze in der Phraseologieforschung / New theoretical and methodological approaches to phraseology. (= Linguistik online 2/06). Internet: http://www.linguistik-online.de/27_06/ (Stand: Mai 2008).
- Haß, Ulrike (Hg.) (2005): Grundfragen der elektronischen Lexikographie. elexiko – das Online-Informationssystem zum deutschen Wortschatz. Berlin/New York.
- Projekt Methoden der Korpusanalyse und -erschließung. Institut für Deutsche Sprache. Mannheim. Homepage. Internet: <http://www.ids-mannheim.de/kl/projekte/methoden> (Stand: Juli 2008).
- Steyer, Kathrin (2000): Usuelle Wortverbindungen des Deutschen. Linguistisches Konzept und lexikografische Möglichkeiten. In: Deutsche Sprache 2/2000, S. 101-125.
- Steyer, Kathrin (2004a): Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven. In: Steyer, Kathrin (Hg.): Wortverbindungen – mehr oder weniger fest. Jahrbuch des Instituts für Deutsche Sprache 2003. Berlin/New York, S. 87-116.
- Steyer, Kathrin (2004b): Idiomatik hypermedial. Zur Repräsentation von Wortverbindungen im Informationssystem Wissen über Wörter. In: Palm-Meister, Christine (Hg.): EUROPHRAS 2000. Internationale Tagung zur Phraseologie vom 15. bis 18. Juni 2000 in Aske/Schweden. (= Stauffenburg Linguistik 25). Tübingen, S. 465-479.
- Steyer, Kathrin (i. Dr.): Zwischen theoretischer Modellierung und praxisnaher Anwendung. Zur korpusgesteuerten Beschreibung usueller Wortverbindungen. In: Mellado Blanco, Carmen (Hg.): Theorie und Praxis der Erstellung von idiomatischen Wörterbüchern. Tübingen.
- Steyer, Kathrin/Lauer, Meike (2006): Elektronische Mehrwortartikel und lexikografische Konzeption. Mannheim. Internet: <http://hypermedia.ids-mannheim.de/elexiko/ModulMV/Konzeption.html> (Stand: Mai 2008).
- Steyer, Kathrin/Lauer, Meike (2007): „Corpus-Driven“: Linguistische Interpretation von Kookkurrenzbeziehungen. In: Kämper, Heidrun/Eichinger, Ludwig M. (Hg.): Sprach-Perspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache. (= Studien zur deutschen Sprache 40). Tübingen, S. 493-509.

5. Referenzwörterbücher in den Mehrwort-Artikeln

Duden 11 (2002) = Duden – Redewendungen. Wörterbuch der deutschen Idiomatik. Bearbeitet von Brigitte Alsleben und Werner Scholze-Stubenrecht. 2., neu bearb. u. aktual. Aufl. Mannheim/Leipzig/Wien/Zürich.

Duden 12 (2002) = Duden – Zitate und Aussprüche. Bearbeitet von Maria Grazia Chiaro und Werner Scholze-Stubenrecht. 2., neu bearb. u. aktual. Aufl. Mannheim/Leipzig/Wien/Zürich.

Röhrich, Lutz (2000): Lexikon der sprichwörtlichen Redensarten. (= Digitale Bibliothek 42). Berlin.

